Short Read Sequencing Analysis Workshop

Day 2

Learning the Linux Compute Environment

Get Github open to srworkshop
Get terminal open

Review of Day 2 Videos

- **Video 1 – SSH and VPN Introduction**
  - How SSH and VPNs work when accessing remote servers
- **Video 2 – Remote Rsync / Reading Files**
  - Looking at files on a remote server and getting them to your computer
- **Video 3 – Searching / editing Files, Pipes, and Outputs**
  - Pieces of your basic toolkit for working on a Linux server
- **Video 4 – Directory Permissions**
  - Managing who can see and execute different things on the server

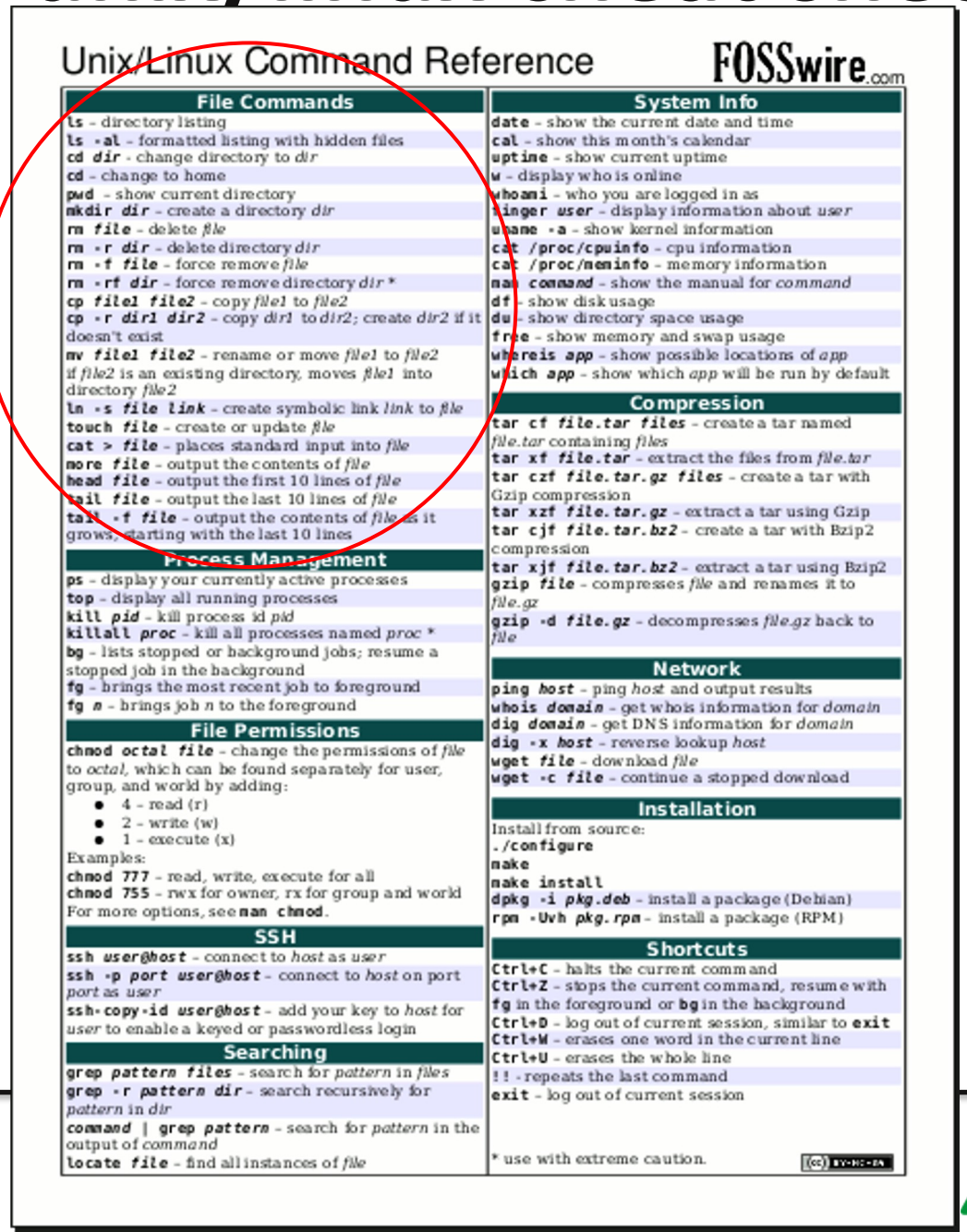
Important things to know about Linux

1. It seems harder but it has advantages
2. File systems (how do I get places?)

Find a unix/linux cheat sheet

The commands might be named weirdly

Cheat sheets on [Github]



Unix/Linux Command Reference

FOSSwire.com

File Commands

- `ls` – directory listing
- `ls -al` – formatted listing with hidden files
- `cd dir` – change directory to *dir*
- `cd` – change to home
- `pwd` – show current directory
- `mkdir dir` – create a directory *dir*
- `rm file` – delete *file*
- `rm -r dir` – delete directory *dir*
- `rm -f file` – force remove *file*
- `rm -rf dir` – force remove directory *dir* *
- `cp file1 file2` – copy *file1* to *file2*
- `cp -r dir1 dir2` – copy *dir1* to *dir2*; create *dir2* if it doesn't exist
- `mv file1 file2` – rename or move *file1* to *file2* if *file2* is an existing directory, moves *file1* into directory *file2*
- `ln -s file link` – create symbolic link *link* to *file*
- `touch file` – create or update *file*
- `cat > file` – places standard input into *file*
- `more file` – output the contents of *file*
- `head file` – output the first 10 lines of *file*
- `tail file` – output the last 10 lines of *file*
- `tail -f file` – output the contents of *file* as it grows, starting with the last 10 lines

Process Management

- `ps` – display your currently active processes
- `top` – display all running processes
- `kill pid` – kill process id *pid*
- `killall proc` – kill all processes named *proc* *
- `bg` – lists stopped or background jobs; resume a stopped job in the background
- `fg` – brings the most recent job to foreground
- `fg n` – brings job *n* to the foreground

File Permissions

`chmod octal file` – change the permissions of *file* to *octal*, which can be found separately for user, group, and world by adding:

- 4 – read (r)
- 2 – write (w)
- 1 – execute (x)

Examples:

- `chmod 777` – read, write, execute for all
- `chmod 755` – rwx for owner, rx for group and world

For more options, see `man chmod`.

SSH

- `ssh user@host` – connect to *host* as *user*
- `ssh -p port user@host` – connect to *host* on port *port* as *user*
- `ssh-copy-id user@host` – add your key to *host* for *user* to enable a keyed or passwordless login

Searching

- `grep pattern files` – search for *pattern* in *files*
- `grep -r pattern dir` – search recursively for *pattern* in *dir*
- `command | grep pattern` – search for *pattern* in the output of *command*
- `locate file` – find all instances of *file*

System Info

- `date` – show the current date and time
- `cal` – show this month's calendar
- `uptime` – show current uptime
- `w` – display who is online
- `whoami` – who you are logged in as
- `finger user` – display information about *user*
- `uname -a` – show kernel information
- `cat /proc/cpuinfo` – cpu information
- `cat /proc/meminfo` – memory information
- `man command` – show the manual for *command*
- `df` – show disk usage
- `du` – show directory space usage
- `free` – show memory and swap usage
- `whereis app` – show possible locations of *app*
- `which app` – show which *app* will be run by default

Compression

- `tar cf file.tar files` – create a tar named *file.tar* containing *files*
- `tar xf file.tar` – extract the files from *file.tar*
- `tar czf file.tar.gz files` – create a tar with Gzip compression
- `tar xzf file.tar.gz` – extract a tar using Gzip
- `tar cjf file.tar.bz2` – create a tar with Bzip2 compression
- `tar xjf file.tar.bz2` – extract a tar using Bzip2
- `gzip file` – compresses *file* and renames it to *file.gz*
- `gzip -d file.gz` – decompresses *file.gz* back to *file*

Network

- `ping host` – ping *host* and output results
- `whois domain` – get whois information for *domain*
- `dig domain` – get DNS information for *domain*
- `dig -x host` – reverse lookup *host*
- `wget file` – download *file*
- `wget -c file` – continue a stopped download

Installation

Install from source:

```
./configure
make
make install
```

- `dpkg -i pkg.deb` – install a package (Debian)
- `rpm -Uvh pkg.rpm` – install a package (RPM)

Shortcuts

- `Ctrl+C` – halts the current command
- `Ctrl+Z` – stops the current command, resume with `fg` in the foreground or `bg` in the background
- `Ctrl+D` – log out of current session, similar to `exit`
- `Ctrl+W` – erases one word in the current line
- `Ctrl+U` – erases the whole line
- `!!` – repeats the last command
- `exit` – log out of current session

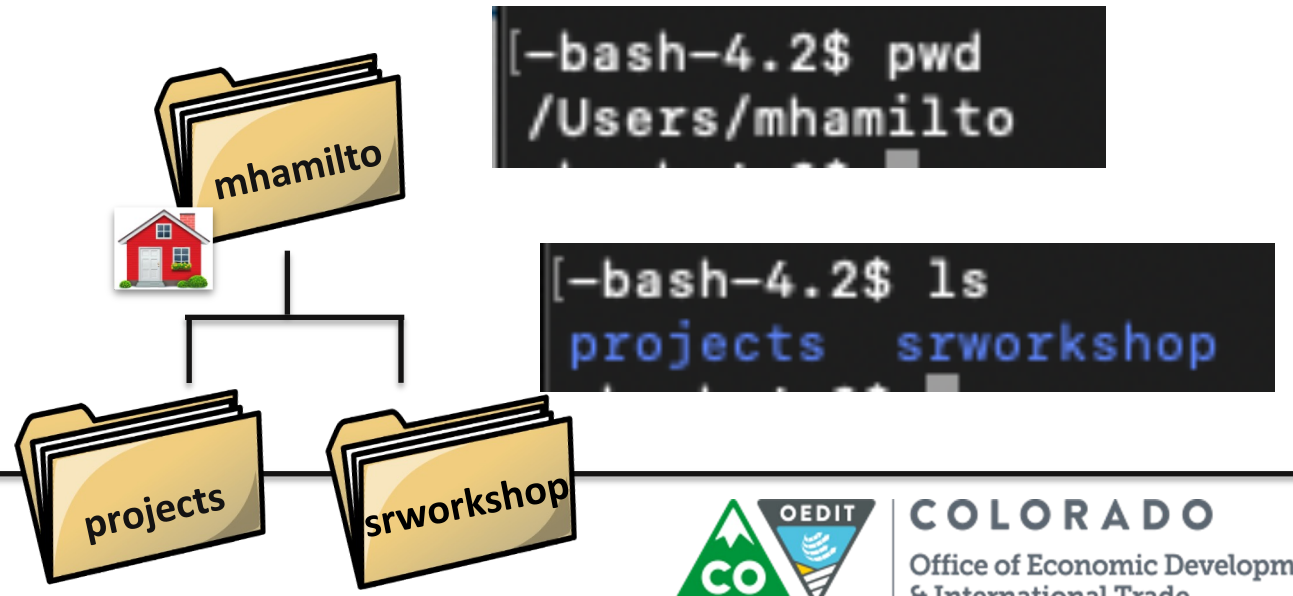
* use with extreme caution.

Common Linux commands

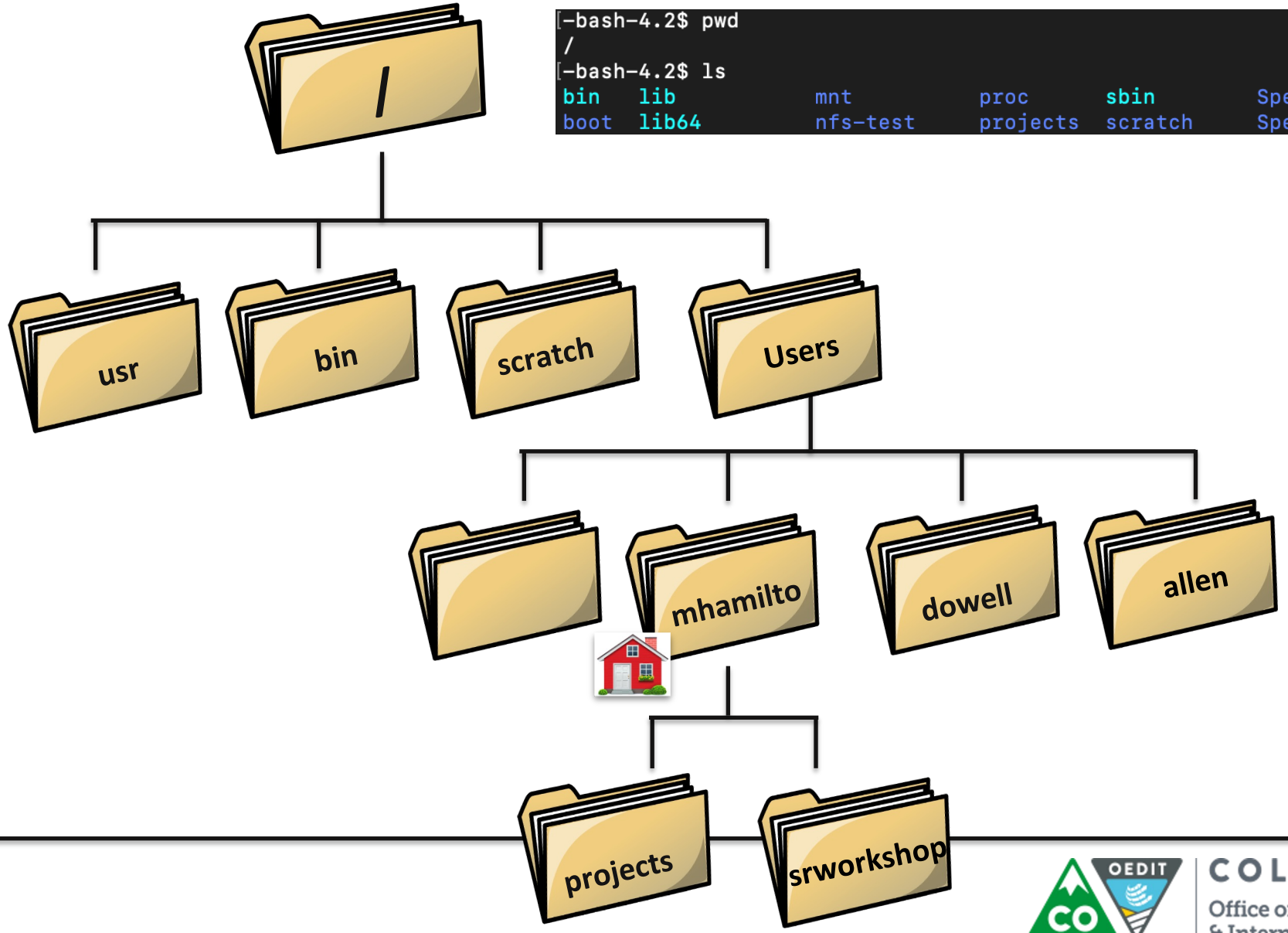
`pwd` “print working directory”

`ls` “list” – list files

`cd` “change directory”

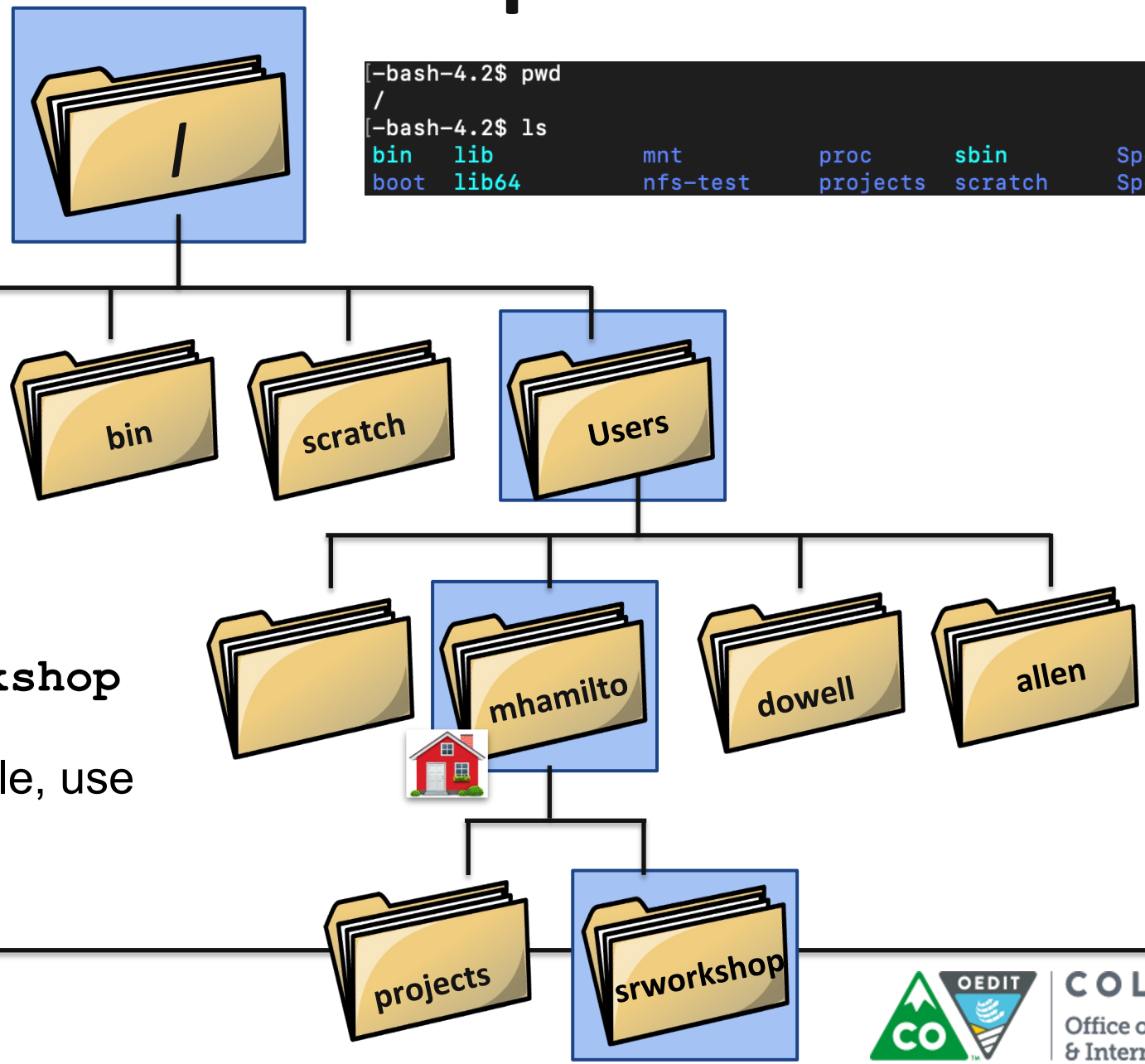


In linux everything is organized as a hierarchy




```
[-bash-4.2$ pwd
/
[-bash-4.2$ ls
bin   lib          mnt          proc      sbin      Spencer_test_1  Users
boot  lib64        nfs-test     projects  scratch   Spencer_test_2  usr
```

Absolute path

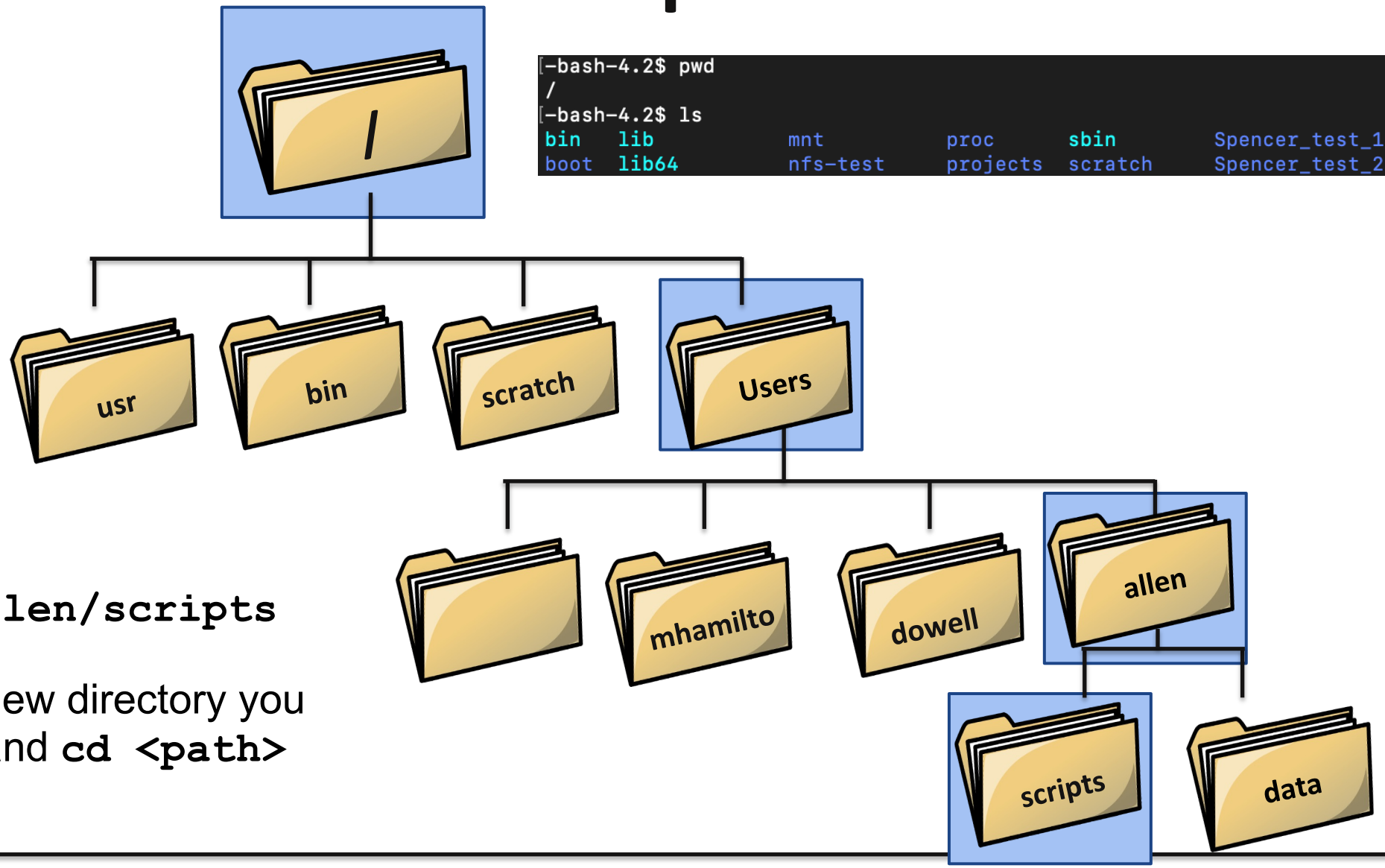



```
[-bash-4.2$ pwd
/
[-bash-4.2$ ls
bin   lib          mnt          proc      sbin      Spencer_test_1  Users
boot  lib64        nfs-test     projects  scratch   Spencer_test_2  usr
```

`/Users/mhamilto/srworkshop`

*To get the absolute path of file, use command **`realpath`**

Absolute path

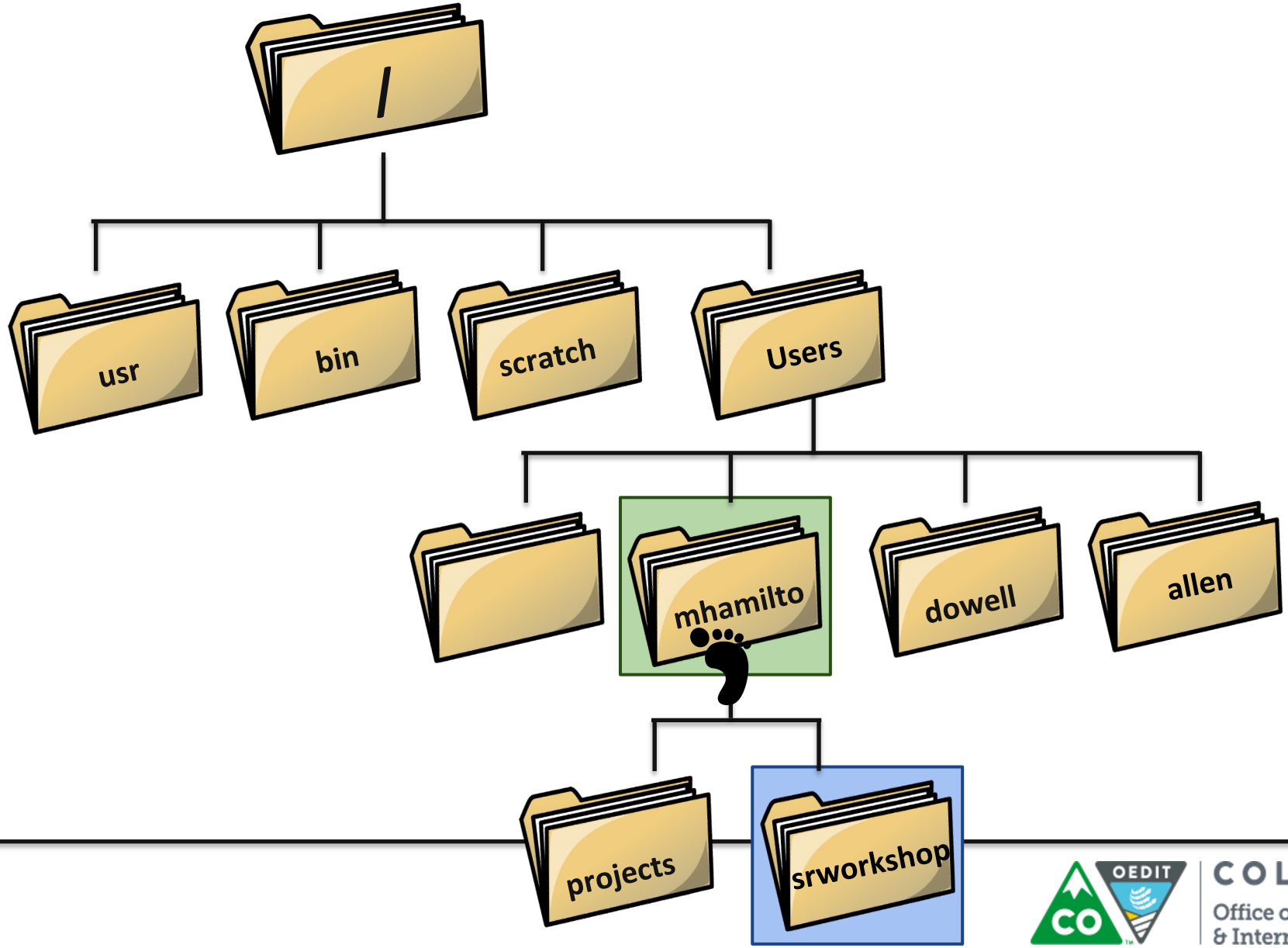



```
[-bash-4.2$ pwd
/
[-bash-4.2$ ls
bin   lib          mnt          proc      sbin      Spencer_test_1  Users
boot  lib64        nfs-test     projects  scratch   Spencer_test_2  usr
```

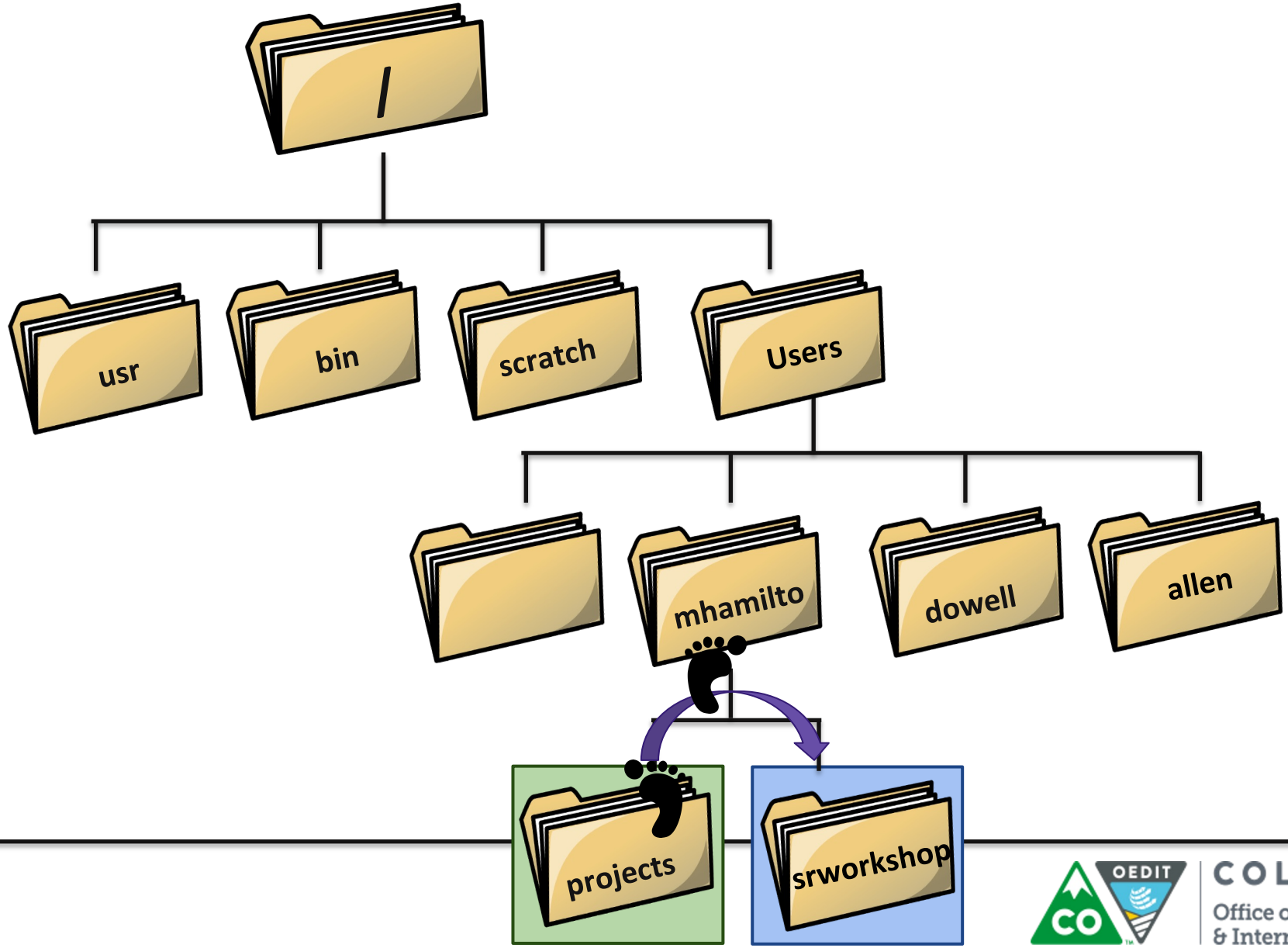
`cd /Users/allen/scripts`

*to move into a new directory you can use command **`cd <path>`**

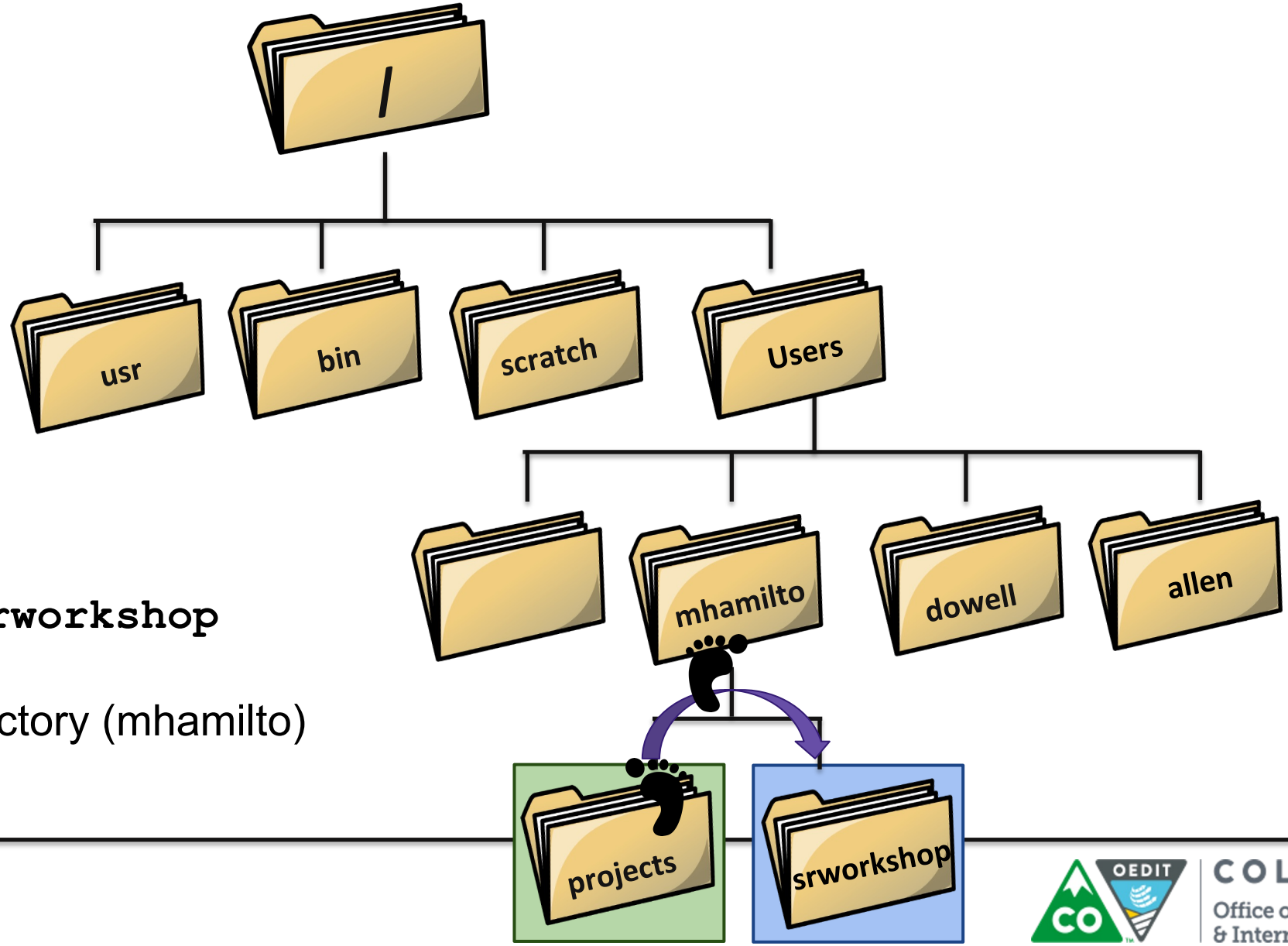
Relative path



Relative path



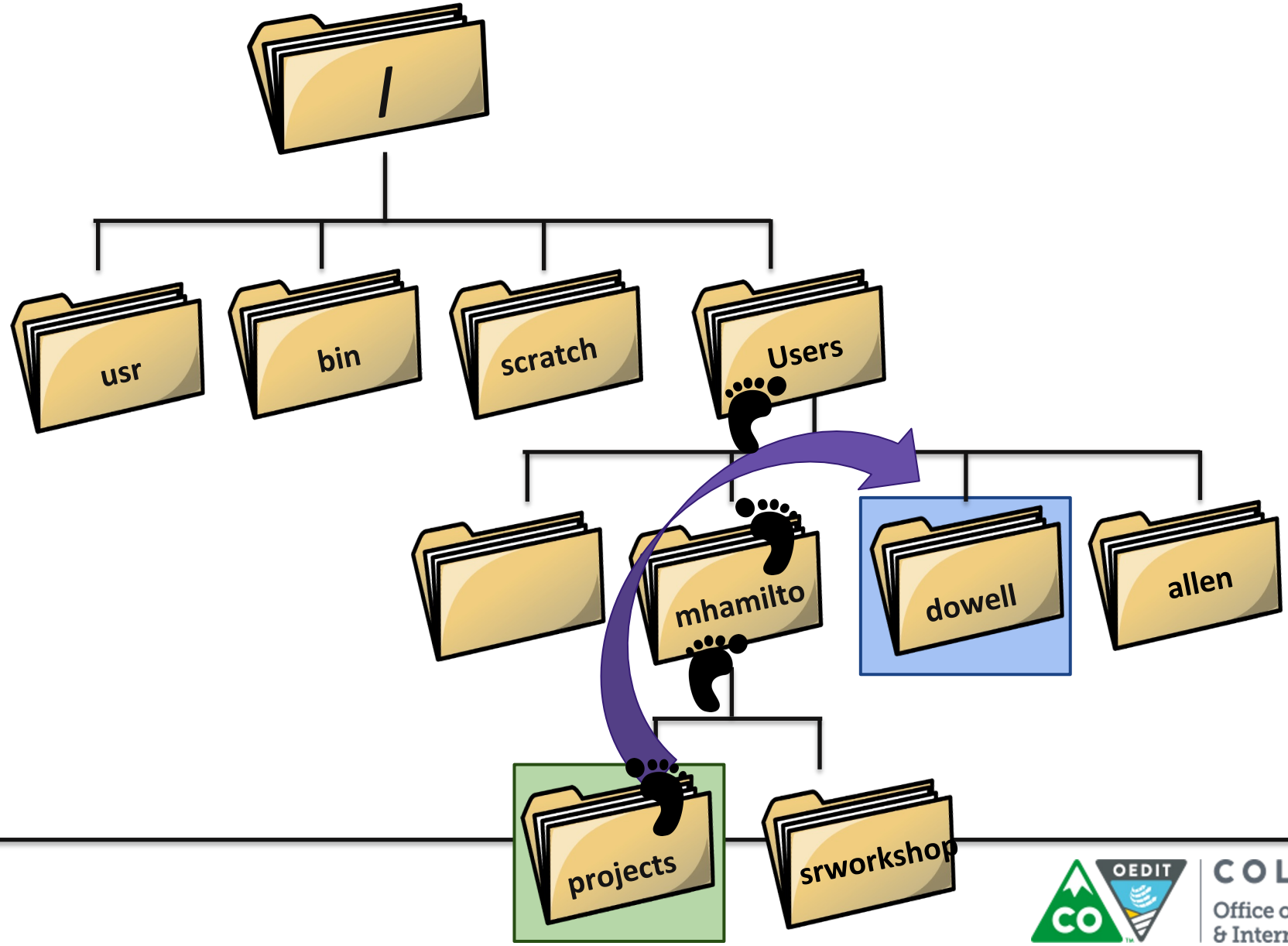
Relative path



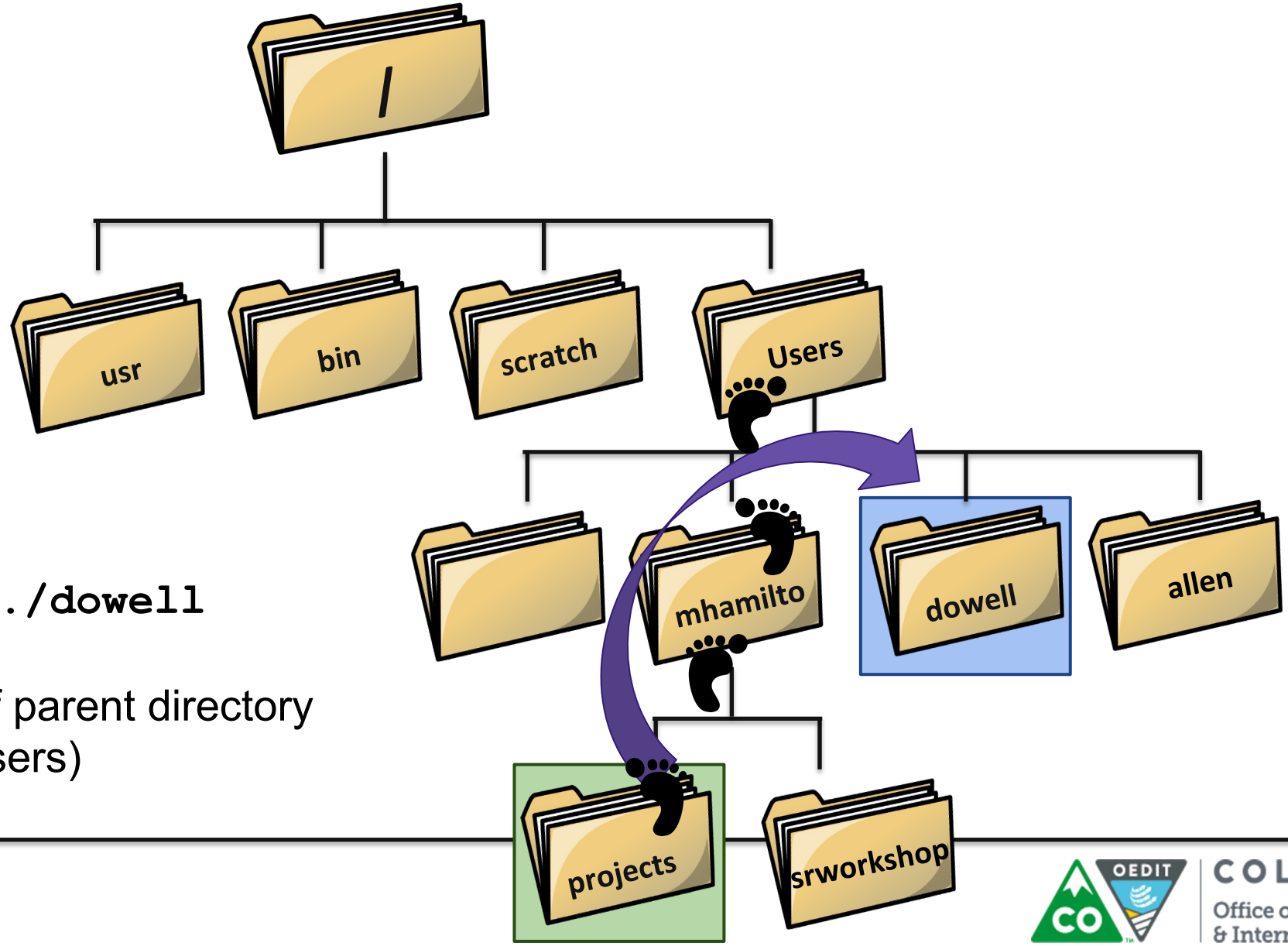

```
cd ../srworkshop
```

.. = parent directory (mhamilto)

Relative path

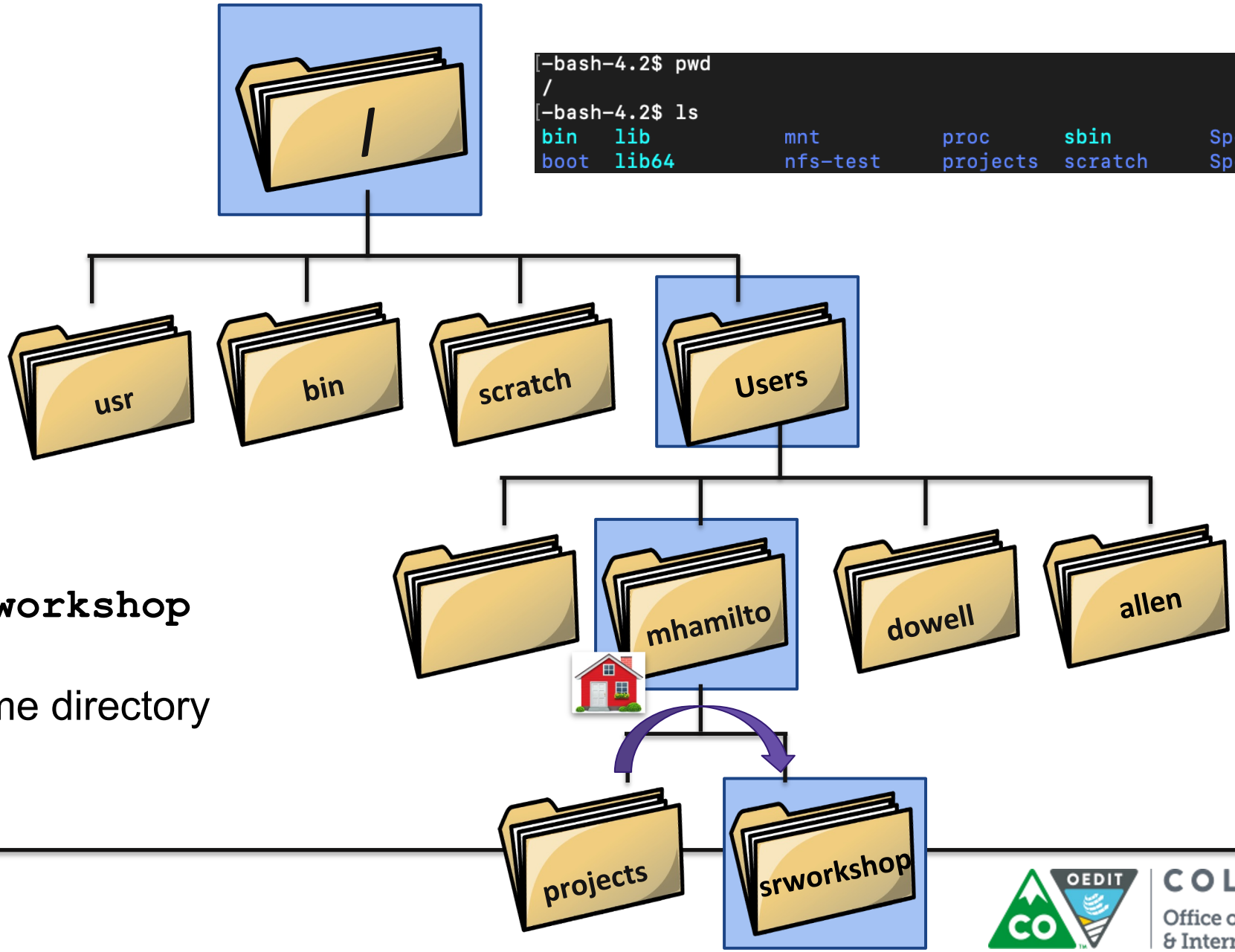


Relative path




```
cd ../../dowell
```

../../ = parent of parent directory (Users)

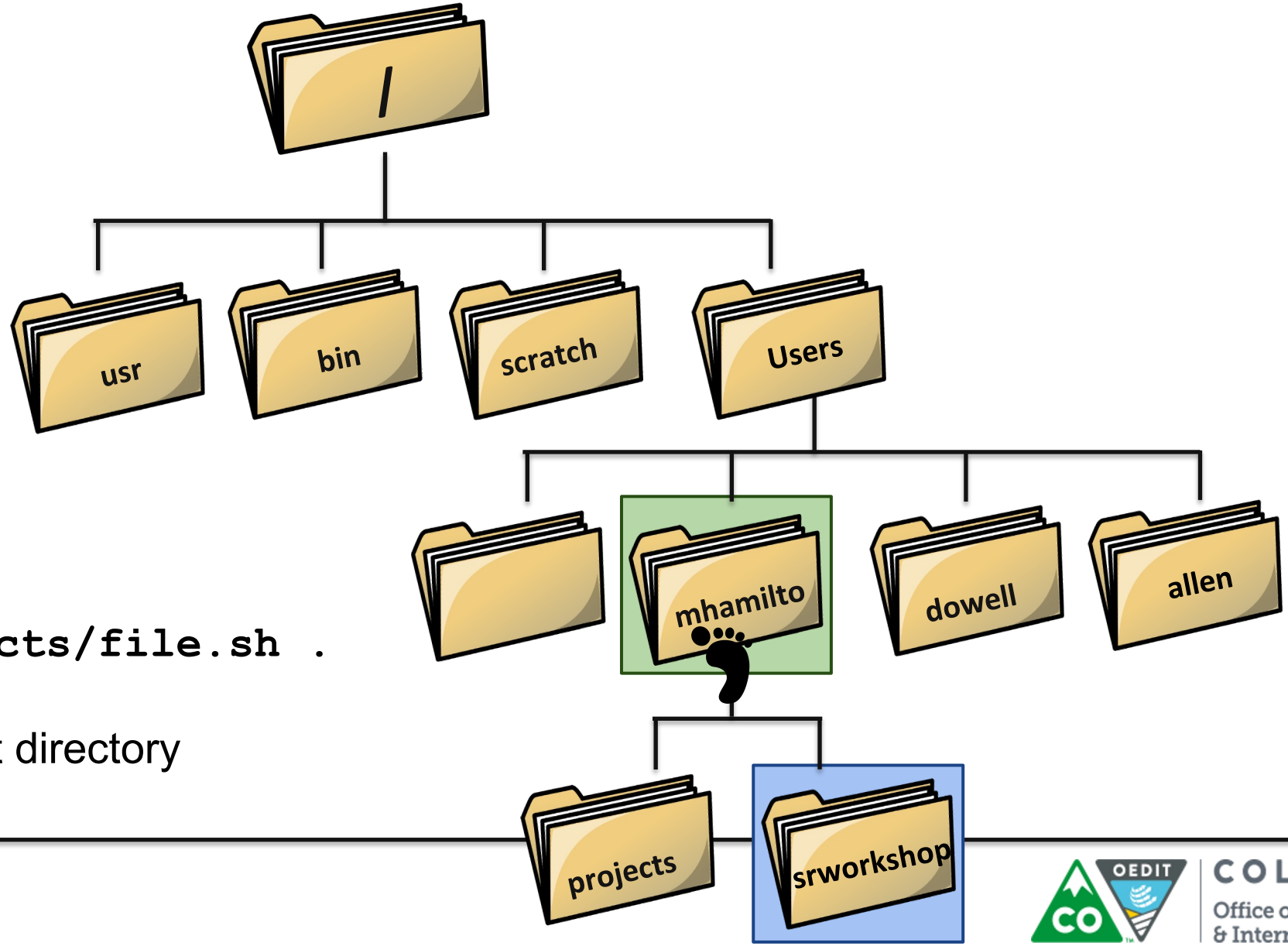



```
[-bash-4.2$ pwd
/
[-bash-4.2$ ls
bin   lib          mnt        proc      sbin      Spencer_test_1  Users
boot  lib64        nfs-test   projects  scratch   Spencer_test_2  usr
```

`cd ~/srworkshop`

~ = your home directory

Relative path




```
rsync ~/projects/file.sh .
```

. = current directory

Questions?

Important things to know about Linux

1. It seems harder but is sometimes WAY easier
2. File systems (how do I get places?)
3. Linux tools don't care about file extensions: *Make sure you are using the correct one.*

Important things to know about Linux

1. It seems harder but is sometimes WAY easier
2. File systems (how do I get places?)
3. Linux tools don't care about file extensions: *Make sure you are using the correct one.*
4. **`man`** command

1. Type **`ls`** in your directory

```
(base) hopekirby@cu-biot-3-10 Demo % ls
R_script.R                  text_file.genes.out
text_file                   text_file.txt
```

2. What else can **`ls`** do?

`man ls`

Highlights

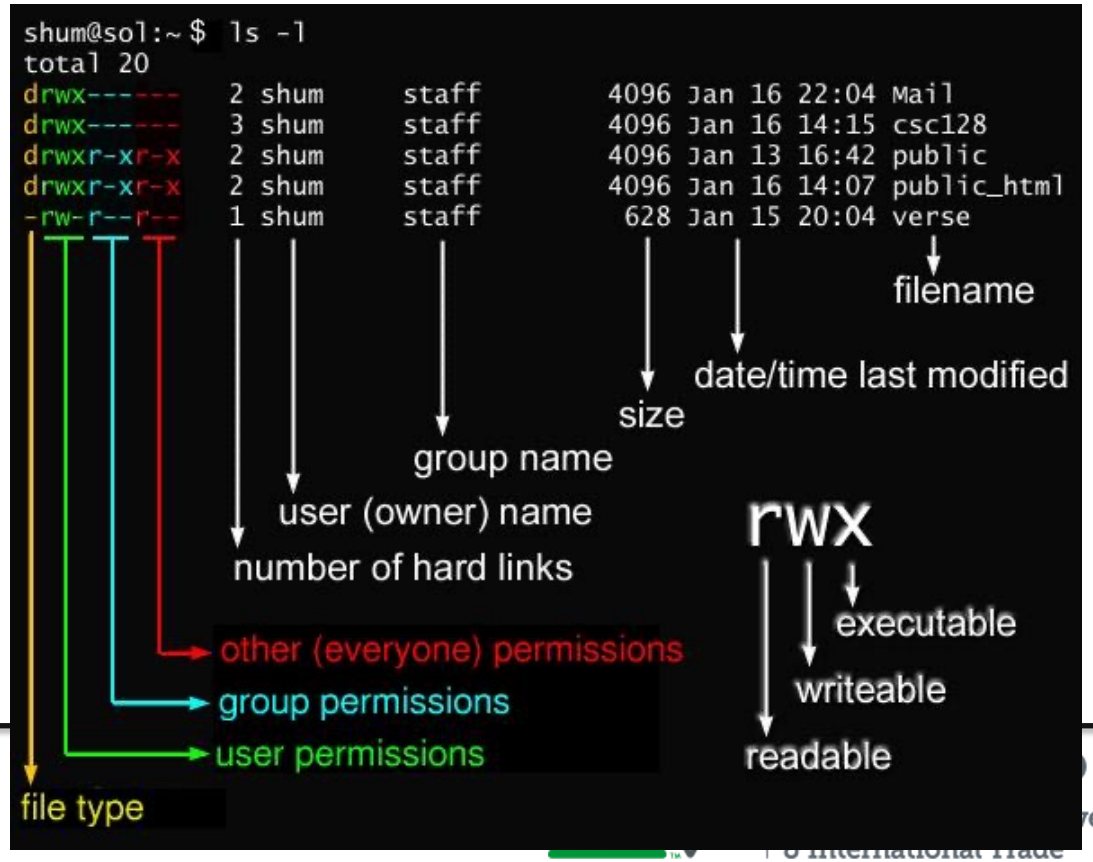
```
-a      Include directory entries whose names begin with a dot ('.').
```

All

```
-l      (The lowercase letter "ell".) List files in the long format, as
        described in the The Long Format subsection below.
```

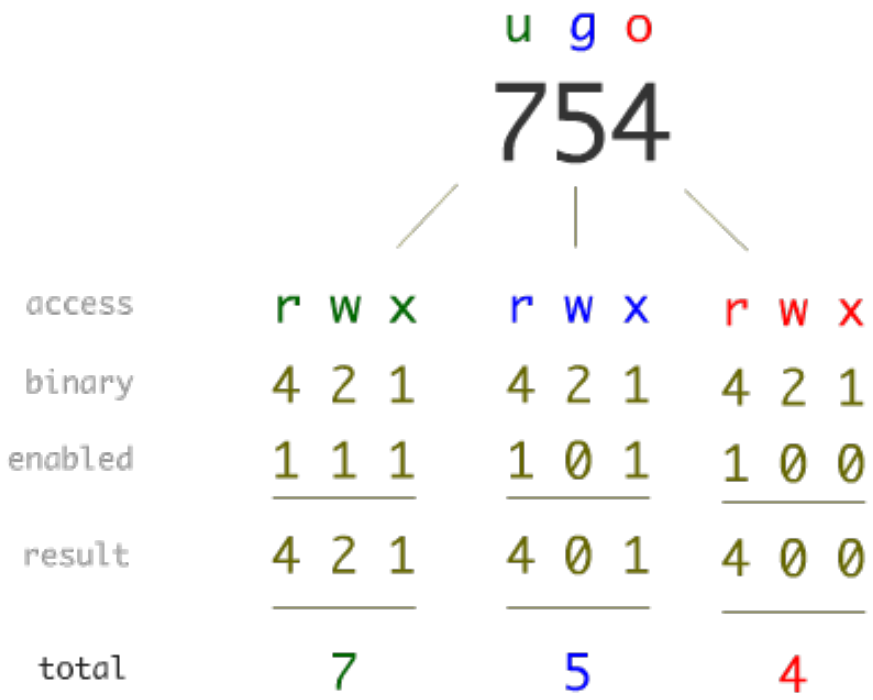
Long

```
The Long Format
  If the -l option is given, the following information is displayed for
  each file: file mode, number of links, owner name, group name, number of
  bytes in the file, abbreviated month, day-of-month file was last
  modified, hour file last modified, minute file last modified, and the
  pathname.  If the file or directory has extended attributes, the
```



File permissions

- Important for sharing files
- Change with **chmod**
- Watch the video!




```
The next three fields are three characters each: owner permissions, group
permissions, and other permissions.  Each field has three character
positions:

     1.   If r, the file is readable; if -, it is not readable.

     2.   If w, the file is writable; if -, it is not writable.

     3.   The first of the following that applies:

               S     If in the owner permissions, the file is not
                     executable and set-user-ID mode is set.  If in the
                     group permissions, the file is not executable and
                     set-group-ID mode is set.

               s     If in the owner permissions, the file is
                     executable and set-user-ID mode is set.  If in the
                     group permissions, the file is executable and
                     setgroup-ID mode is set.

               x     The file is executable or the directory is
                     searchable.

               -     The file is neither readable, writable,
                     executable, nor set-user-ID nor set-group-ID mode,
                     nor sticky.  (See below.)

          These next two apply only to the third character in the last
          group (other permissions).

               T     The sticky bit is set (mode 1000), but not execute
                     or search permission.  (See chmod(1) or
                     sticky(7).)

               t     The sticky bit is set (mode 1000), and is
                     searchable or executable.  (See chmod(1) or
                     sticky(7).)
```

Highlights cont…

`ls -lah`

Highlights:

```
-a      Include directory entries whose names begin with a dot ('.').
```

All

```
-l      (The lowercase letter "ell".) List files in the long format, as
        described in the The Long Format subsection below.
```

```
The Long Format
  If the -l option is given, the following information is displayed for
  each file: file mode, number of links, owner name, group name, number of
  bytes in the file, abbreviated month, day-of-month file was last
  modified, hour file last modified, minute file last modified, and the
  pathname.  If the file or directory has extended attributes, the
```

Long

```
-h      When used with the -l option, use unit suffixes: Byte, Kilobyte,
        Megabyte, Gigabyte, Terabyte and Petabyte in order to reduce the
        number of digits to four or fewer using base 2 for sizes.  This
        option is not defined in IEEE Std 1003.1-2008 ("POSIX.1").
```

"Human readable"

Questions?

Start Worksheet Part 1

Remember to make the terminal comfortable to work in:

- Make the windows large
- Open multiple windows and/or tabs
- Change font sizes etc. ((For Mac: Command + Plus/Minus, For Windows, find Text size in Preferences)

You can move on to Part 2 when finished with Part 1

If you haven’t already, start Worksheet Part 2

Done with Part 2 already?

1. Get started on the homework!
2. Practice the commands in the cheat sheets linked on Github
3. Help a peer out!

Questions?

Don’t forget the homework

Help session in JSCBB A108 from 1-3pm

Watch videos for Day 3